

LEARNING AND USING GENERALIZED STRING PATTERNS FOR INFORMATION EXTRACTION

ABSTRACT OF THE DISCLOSURE

5 The present invention relates to extracting
information from an information source. During
extraction, strings in the information source are
accessed. These strings in the information source are
matched with generalized extraction patterns that
10 include words and wildcards. The wildcards denote
that at least one word in an individual string can be
skipped in order to match the individual string to an
individual generalized extraction pattern.